Multi Users information theory Semester A 2009/10

Homework Set #5

1) **MAC with common information**. Consider a DM-MAC $P_{Y|X_1,X_2}$ with three independent uniformly distributed messages $W_0 \in [1,...,2^{nR0}]$, $W_1 \in [1,...,2^{nR1}]$, and $W_2 \in [1,...,2^{nR2}]$. The first encoder maps each pair (w_0,w_1) into a codeword $x_1^n(w_0,w_1)$ and the second maps each pair (w_0,w_2) into a codeword $x_2^n(w_0,w_2)$. The decoder upon receiving y^n, finds an estimate $(\hat{w}_0,\hat{w}_1,\hat{w}_2)$ of the messages sent. The probability of decoding error is:

$$P_e^{(n)} = \Pr\left((\hat{W}_0,\hat{W}_1,\hat{W}_2) \neq (W_0,W_1,W_2)\right). \tag{1}$$

Show that the capacity region for this channel is given by the set of rate triples (R_0,R_1,R_2) such that

$$\begin{aligned} R_1 &\leq I(X_1;Y|X_2,U), \\ R_2 &\leq I(X_2;Y|X_1,U), \\ R_1+R_2 &\leq I(X_1,X_2;Y|U), \\ R_0+R_1+R_2 &\leq I(X_1,X_2;Y), \end{aligned} \tag{2}$$

for some $p(u)p(x_1|u)p(x_2|u)$. You need to prove achievability and converse. (Hint: In proving the converse you may use the identification $U_i = W_0$.)

MAC with common information solution:

We show that the capacity is given by the set of rate triples (R_0,R_1,R_2) such that

$$\begin{aligned} R_1 &\leq I(X_1;Y|X_2,U), \\ R_2 &\leq I(X_2;Y|X_1,U), \\ R_1+R_2 &\leq I(X_1,X_2;Y|U), \\ R_0+R_1+R_2 &\leq I(X_1,X2;Y), \end{aligned}$$

for some $p(u)p(x_1|u)p(x_2|u)$. Note that this set is convex and therefore there is no need for further convexification.

Proof of achievability: Fix $p(u)p(x_1|u)p(x_2|u)$. Generate 2^{nR_0} sequences $u^n(w_0)$ according to $p(u^n) = \prod_{i=1}^n p(u_i)$. For each sequence u^n generater 2^{nR_1} sequences $x_1^n(w_0,w_1)$ according to $p(x_1^n|u^n) = \prod_{i=1}^n p(x_{1i}|u_i)$ and 2^{nR_2} sequences $x_2^n(w_0,w_2)$ according to $p(x_2^n|u^n) = \prod_{i=1}^n p(x_{2i}|u_i)$. To transmit (w_0,w_1,w_2). the first transmitter sends $x_1^n(w_0,w_1)$

and the second transmitter sends $x_2^n(w_0, w_2)$.

The decoder upon receiving y^n, looks for the unique $(\hat{w_0}, \hat{w_1}, \hat{w_2})$ such that $\{(u^n(\hat{w_0}), x_1^n(\hat{w_0}\hat{w_1}), x_2^n(\hat{w_0}\hat{w_2}), y^n) \in A_\epsilon^{(n)}\}$.

Probability or error: Assume $(1, 1, 1)$ is sent and define the events

$$E_{ijk} = \{(u^n(i), x_1^n(i,j), x_2^n(i,k), y^n) \in A_\epsilon^{(n)}\}.$$

Then the probability of decoding error is:

$$P_e^{(n)} \le P(E_{111}^c) + \textstyle\sum_{i\neq 1,j,k} P(E_{ijk}) + \sum_{j\neq 1} P(E_{1j1}) + \sum_{k\neq 1} P(E_{11k}) + \sum_{j,k\neq 1} P(E_{1jk}).$$

$P(E_{111}^c) \to 0$ by AEP. Now consider the third term

$$\begin{aligned} \sum_{j\neq 1} P(E_{1j1}) &\le 2^{nR_1} \sum_{(u^n,x_1^n,x_2^n,y^n)\in A_\epsilon^{(n)}} p(u^n)p(x_1^n|u^n)p(x_2^n|u^n)p(y^n|x_2^n,u^n) \\ &\overset{(a)}{\le} 2^{n(R_1 - I(X_1;Y|X_2,U)+7\epsilon)}. \end{aligned}$$

where

(a) follows from the jointly typical set, when u and x_2 are known.

Thus if $R_1 < I(X_1; Y|X_2, U) - 7\epsilon$, the third term in the bound on $P_e^{(n)}$ approaches 0 as $n \to \infty$. The fourth term follows similarly and we obtain the requirement that $R_2 \le I(X_2; Y|X_1, U) - 7\epsilon$. The last term approaches 0 as $n \to \infty$ if $R_1 + R_2 < I(X_1, X_2; Y|U) - 7\epsilon$. Finally consider the second term

$$\begin{aligned} \sum_{i\neq 1,j,k} P(E_{ijk}) &\le 2^{n(R_0+R_1+R_2)} \sum_{(u^n,x_1^n,x_2^n,y^n)\in A_\epsilon^{(n)}} p(u^n)p(x_1^n|u^n)p(x_2^n|u^n)p(y^n) \\ &\le 2^{n(R_0+R_1+R_2)} 2^{n(H(U,X_1,X_2,Y)+\epsilon)} 2^{-n(H(U)-\epsilon)} 2^{-n(H(X_1|U)-2\epsilon)} 2^{-n(H(X_2|U)-2\epsilon)} 2^{-n(H(Y)-\epsilon)}. \end{aligned}$$

Thus if

$$\begin{aligned} R_0 + R_1 + R_2 &< H(U, X_1, X_2, Y) - H(U) - H(X_1|U) - H(X_2|U) - H(Y) - 7\epsilon \\ &= I(X_1, X_2; Y) - 7\epsilon, \end{aligned}$$

the second sum of the bound $\Pr_e^{(n)}$ goes to 0 as $n \to \infty$.

Proof of converse:

First consider

$$\begin{aligned} nR_1 &= H(W_1) \\ &= H(W_1|W_0, W_2) \\ &\le I(W_1; Y^n|W_0, W_2) + n\epsilon_n \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} I(W_1; Y_i | W_0, W_2, Y^{i-1}) + n\epsilon_n \\
&\leq \sum_{i=1}^{n} H(Y_i|W_0, W_2) - H(Y_i|W_0, W_1, W_2, Y^{i-1}) + n\epsilon_n \\
&\leq \sum_{i=1}^{n} H(Y_i|X_{2i}, W_0, W_2) - H(Y_i|X_{1i}, X_{2i}, W_0, W_1, W_2, Y^{i-1}) + n\epsilon_n \\
&\leq \sum_{i=1}^{n} H(Y_i|W_0, X_{2i}) - H(Y_i|W_0, X_{1i}, X_{2i}) + n\epsilon_n \\
&= \sum_{i=1}^{n} I(X_{1i}; Y_i | X_{2i}, U_i) + n\epsilon_n,
\end{aligned}
$$

where $U_i = W_0$, so $p(u_i, x_{1i}, x_{2i}) = p(u_i)p(x_{1i}|u_i)p(x_{2i}|u_i)$. Similarly it can be shown that $nR_2 \leq \sum_{i=1}^{n} I(X_{2i}; Y_i|X_{1i}, U_i) + n\epsilon_n$ and $n(R_1 + R_2) \leq \sum_{i=1}^{n} I(X_{1i}, X_{2i}; Y_i|U_i) + n\epsilon_n$.

It is also easy to show that $n(R_0 + R_1 + R_2) \leq \sum_{i=1}^{n} I(X_{1i}, X_{2i}; Y_i) + n\epsilon_n$. Now, introducing a time-sharing random variable Q, we have

$$
\begin{aligned}
R_1 &\leq I(X_{1Q}; Y_Q | X_{2Q}, U_Q, Q) = I(X_1; Y | X_2, U) + \epsilon, \\
R_1 &\leq I(X_{2Q}; Y_Q | X_{1Q}, U_Q, Q) = I(X_2; Y | X_1, U) + \epsilon, \\
R_1 + R_2 &\leq I(X_{1Q}, X_{2Q}; Y_Q | U_Q, Q) = I(X_1, X_2; Y | U) + \epsilon, \\
R_0 + R_1 + R_2 &\leq I(X_{1Q}, X_{2Q}; Y_Q | Q) = I(X_1, X2; Y) + \epsilon,
\end{aligned}
$$

where $U = (U_Q, Q),\ X_1 = X_{1Q},\ X_2 = X_{2Q}$ and $Y = Y_Q$, and

$$p(u, x_1, x_2, y) = p(u)p(x_1|u)p(x_2|u)p(y|x_1, x_2).$$

2) **Strong ϵ-typicality**. Achievability proofs involving *covering*, e.g., for the rate distortion theorem, require that we find a good lower bound on the probability that one specific typical sequence x^n is jointly typical with a randomly drawn sequence Y^n. Using strong typicality, the desired lower bound can be established. Let (X_i, Y_i) be drawn i.i.d. $\sim P(x, y)$ and assume that the cardinalities $\mathcal{X}, \mathcal{Y}$ are finite. Let the marginals of X and Y be $P(x)$ and $P(y)$, respectively (you may use ideas and results from methods of types to solve the exercise). We use in this exercise a specific notation $\delta(\epsilon)$ that implies that $\delta(\epsilon) \to 0$ as $\epsilon \to 0$.

   a) Show that if $x^n \in T_\epsilon^{(n)}(X)$, then

$$p(x^n) \doteq 2^{n(H(X) \pm \delta(\epsilon))}. \tag{3}$$

   b) Show that $\Pr(X^n \in T_\epsilon^{(n)}(X)) \to 1$, as $n \to \infty$.

   c) Show that

$$|T_\epsilon^{(n)}(X)| \doteq 2^{n(H(X) \pm \delta(\epsilon)}. \tag{4}$$

d) Let $x^n \in T_\epsilon^{(n)}(X)$, and let $T_\epsilon^{(n)}(Y|x^n)$ be the set of y^n sequences such that $(x^n, y^n) \in T_\epsilon^{(n)}(X,Y)$. Show that

$$|T_\epsilon^{(n)}(Y|x^n)| \doteq 2^{n(H(Y|X)\pm\delta(\epsilon))}. \tag{5}$$

e) Let $x^n \in T_\epsilon^{(n)}(X)$, and Y^n be drawn independently of x^n i.i.d. $\sim P(y)$. Show that

$$\Pr(x^n, Y^n \in T_\epsilon^{(n)}(X,Y)| \doteq 2^{n(I(X,Y)\pm\delta(\epsilon)}. \tag{6}$$

(Note that in (d) and (e) the bounds do not depend on x^n.)

Strong ϵ-typicality solution:

a) We recall that $T_\epsilon^n(X) = \{x^n : |P_{x^n}(a) - p_X(a)| < \epsilon, p_X(a) = 0 \Rightarrow P_{x^n}(a) = 0\}$.
log is a continues function thus $|x-y| < \epsilon \to |\log(x) - \log(y)| = |\log \frac{x}{y}| < \delta(\epsilon)$, and so we obtain-

$$\begin{aligned} D(P_{X^n}||p(x)) &= \sum_{x\in\mathcal{X}} P_{X^n} \log \frac{P_{X^n}}{p(x)} \\ &\overset{(a)}{<} \sum_{x\in\mathcal{X}} P_{X^n}\delta(\epsilon) \\ &= |\mathcal{X}|\delta(\epsilon) = \delta \end{aligned}$$

where

(a) follows from the continuity of the log function, as explained above.

From method of types, we have- $Q^n(x^n) \doteq 2^{-n(D(p||q)+H(p))}$, thus for $x^n \in T_\epsilon^n(X)$, we have-

$$p(x^n) = 2^{-n(D(P_{x^n}||p(x))+H(X))} \geq 2^{-n(H(X)+\epsilon)}.$$

It is also clear that-

$$p(x^n) \leq 2^{-n(H(X)-\epsilon)},$$

thus we obtain $p(x^n) \doteq 2^{-n(H(X)\pm\epsilon)}$.

b) We consider-

$$\begin{aligned} 1 - \Pr(x^n \in T_\epsilon^n(X)) &= \Pr(x^n \notin T_\epsilon^n(X)) \\ &= \sum_{p(x^n):D(P_{x^n}||p(x))>\epsilon} p(x^n) \\ &\overset{(a)}{\doteq} \sum_{p(x^n):D(P_{x^n}||p(x))>\epsilon} 2^{-n(H(X)+\epsilon)} \\ &\leq |\mathbb{P}_n| 2^{-n(H(X)+\epsilon)} \\ &\overset{(b)}{\leq} (n+1)^{|\mathcal{X}|} 2^{-n(H(X)+\epsilon)} \end{aligned}$$

$$= \quad 2^{-n(H(X)+\epsilon-\frac{\log(n+1)}{n}|\mathcal{X}|)} \to 0.$$

where

(a) follows from strong ϵ typicality.

(b) follows from the fact that there are, as an upper bound, $(n+1)^{|\mathcal{X}|}$ types.

Thus we obtain $\Pr(x^n \in T_\epsilon^n(X)) \to 1$ as $n \to \infty$.

c) Using the result from (b) we have-

$$\begin{aligned}
\Pr(x^n \in T_\epsilon^n(X)) &= \sum_{x^n \in T_\epsilon^n(X)} p(x^n) \\
&\doteq \sum_{x^n \in T_\epsilon^n(X)} 2^{-n(H(X)\pm\epsilon)} \\
&= |T_\epsilon^n(X)| 2^{-n(H(X)\pm\epsilon)}.
\end{aligned}$$

Note that if $\lim a_n = 1 \Rightarrow a_n \doteq 1$, thus-

$$\begin{aligned}
1 &\doteq \Pr(x^n \in T_\epsilon^n(X)) \\
&\doteq |T_\epsilon^n(X)| 2^{-n(H(X)\pm\epsilon)}.
\end{aligned}$$

Thus we obtain $|T_\epsilon^n(X)| \doteq 2^{-n(H(X)\pm\epsilon)}$.

d) Let $x^n \in T_\epsilon^n(X)$, and y^n is drawn$\sim p(y_i|x_i)$, thus (x_i, y_i) is drawn i.i.d $p_{X,Y}$, and from (a) we have $\Pr((x^n, y^n) \in T_\epsilon^n(X,Y)) \to 1$. Thus-

$$\begin{aligned}
1 &\doteq \Pr((x^n, y^n) \in T_\epsilon^n(X,Y)) \\
&\overset{(a)}{=} \sum_{x^n \in T_\epsilon^n(X)} p((x^n, y^n) \in T_\epsilon^n(Y|x^n)) \\
&\overset{(b)}{=} \sum_{x^n \in T_\epsilon^n(X)} \sum_{y^n \in T_\epsilon^n(Y|x^n)} p(x^n, y^n) \\
&\overset{(c)}{\doteq} \sum_{x^n \in T_\epsilon^n(X)} \sum_{y^n \in T_\epsilon^n(Y|x^n)} 2^{-n(H(X,Y)\pm\epsilon)} \\
&= \sum_{x^n \in T_\epsilon^n(X)} |T_\epsilon^n(Y|x^n)| 2^{-n(H(X,Y)\pm\epsilon)} \\
&= |T_\epsilon^n(X)||T_\epsilon^n(Y|x^n)| 2^{-n(H(X,Y)\pm\epsilon)} \\
&\doteq |T_\epsilon^n(Y|x^n)| 2^{-n(H(X,Y)-H(X)\pm\epsilon)}
\end{aligned}$$

where

(a) follows from the definition of $T_\epsilon^n(X,Y)$.

(b) follows from the definition of $T_\epsilon^n(Y|x^n)$.

(c) follows from the fact that $(x^n, y^n) \in T_\epsilon^n(X,Y)$.

Thus we obtain $|T_\epsilon^n(Y|x^n)| \doteq 2^{n(H(X,Y)-H(X)\pm\epsilon)} = 2^{n(H(Y|X)\pm\epsilon)}$.

e) Now we have $x^n \in T_\epsilon^n(X)$, $Y^n \sim p(y)$ i.i.d, independent of x^n.

$$\begin{aligned}
\Pr((x^n, Y^n) \in T_\epsilon^n(X,Y)) &\overset{(a)}{=} \sum_{y^n \in T_\epsilon^n(Y|x^n)} p(y^n) \\
&\overset{(b)}{\doteq} \sum_{y^n \in T_\epsilon^n(Y|x^n)} 2^{-n(H(Y)\pm\epsilon)} \\
&= |T_\epsilon^n(Y|x^n)| 2^{-n(H(Y)\pm\epsilon)} \\
&= 2^{-n(H(Y)\pm\epsilon - H(Y|X)\pm\epsilon')} \\
&= 2^{-n(I(X;Y)\pm\epsilon)}.
\end{aligned}$$

where

(a) follows from the definition of $T_\epsilon^n(Y|x^n)$.

(b) follows from the fact that $y^n \in T_\epsilon^n(Y)$.